





RESEARCH ARTICLE | MARCH 01 2024

## GenIce-core: Efficient algorithm for generation of hydrogen-disordered ice structures

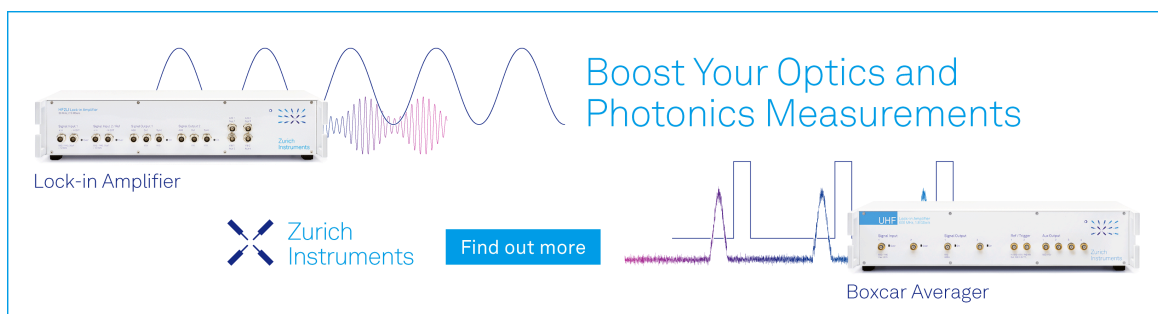
Special Collection: [Modular and Interoperable Software for Chemical Physics](#)

Masakazu Matsumoto   ; Takuma Yagasaki  ; Hideki Tanaka 




*J. Chem. Phys.* 160, 094101 (2024)

<https://doi.org/10.1063/5.0198056>



Boost Your Optics and Photonics Measurements

Lock-in Amplifier

 Zurich Instruments

[Find out more](#)

Boxcar Averager

# GenIce-core: Efficient algorithm for generation of hydrogen-disordered ice structures

Cite as: J. Chem. Phys. 160, 094101 (2024); doi: 10.1063/5.0198056

Submitted: 16 January 2024 • Accepted: 9 February 2024 •

Published Online: 1 March 2024



View Online



Export Citation



CrossMark

Masakazu Matsumoto,<sup>1,a)</sup> Takuma Yagasaki,<sup>2</sup> and Hideki Tanaka<sup>1,3</sup>

## AFFILIATIONS

<sup>1</sup>Research Institute for Interdisciplinary Science, Okayama University, Okayama 700-8530, Japan

<sup>2</sup>Division of Chemical Engineering, Graduate School of Engineering Science, Osaka University, Osaka 560-8531, Japan

<sup>3</sup>Toyota Physical and Chemical Research Institute, Nagakute 480-1192, Japan

**Note:** This paper is part of the JCP Special Topic on Modular and Interoperable Software for Chemical Physics.

**a)** Author to whom correspondence should be addressed: [vitroid@gmail.com](mailto:vitroid@gmail.com)

## ABSTRACT

Ice is different from ordinary crystals because it contains randomness, which means that statistical treatment based on ensemble averaging is essential. Ice structures are constrained by topological rules known as the ice rules, which give them unique anomalous properties. These properties become more apparent when the system size is large. For this reason, there is a need to produce a large number of sufficiently large crystals that are homogeneously random and satisfy the ice rules. We have developed an algorithm to quickly generate ice structures containing ions and defects. This algorithm is provided as an independent software module that can be incorporated into crystal structure generation software. By doing so, it becomes possible to simulate ice crystals on a previously impossible scale.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0198056>

## INTRODUCTION

Ice is ubiquitous and is one of the major constituent solids of the universe.<sup>1</sup> However, ice has a variety of interesting properties that are not commonly found in other materials, e.g., ice floats on liquid water; as many as 20 different crystal structures exist,<sup>2</sup> and it has a negative thermal expansion coefficient at low temperatures and pressures.<sup>3</sup> What makes ice truly unique, however, is hydrogen disorder. In normal ice I, oxygen atoms are located at lattice points, while hydrogen atoms stochastically occupy one of two positions between adjacent pairs of oxygen atoms under some constraints. This randomness is what causes the residual entropy in ice.<sup>4</sup>

The finite residual entropy found in ice indicates the presence of inherent randomness. This randomness is subject to certain constraints known as the ice rules.<sup>5</sup> In all ice phases, including normal ice, every water molecule forms four hydrogen bonds with four neighboring molecules. Two bonds are donated to adjacent two molecules, while the other two are donated by the remaining two neighboring molecules. The mole frac-

tion of point defects that violate the ice rules is typically only  $10^{-6}$ .<sup>6</sup> Therefore, almost all water molecules in ice follow the ice rules.

The ice rules not only govern the local arrangement but also determine various ice-specific properties over long distances. For instance, in proton transfer in ice, protons (hydrogen cations) move rapidly along homodromic chains.<sup>7</sup> This movement is accompanied by inversions of the hydrogen bonds along the path, which means that another proton cannot move in the same direction along the same path.<sup>8</sup> The reactant and the background can be treated separately in a typical chemical reaction. However, in a reaction involving protons in ice, the reaction changes the structure of the background ice. It was also demonstrated that higher-order structures in hydrogen-disordered ice homogenize the binding energies of water molecules in ice, which is a prerequisite for the Pauling entropy approximation.<sup>9,10</sup> The ice rules have a close relationship with ice polarization. The concept of localized nucleation and its growth, as in ordinary nucleation processes, is not well-suited for the phase transition from a ferroelectric hydrogen-ordered ice to a depolarized hydrogen-disordered ice of the same skele-

tal structure (such as ice XI and Ih). The topological constraints of the ice rules force string-like atypical structural changes during the transition.<sup>11</sup>

To accurately evaluate the physical properties of ice by computer simulations, it is necessary to generate a large number of unbiased, random, and hydrogen-disordered structures for a single arrangement of oxygen atoms. Physical properties can be calculated as their average. Rahman and Stillinger were the first to develop an algorithm for generating hydrogen-disordered ice that satisfies the ice rules.<sup>12</sup> This algorithm starts with a structure that satisfies the ice rules but is not random. Randomness is introduced by repeatedly reversing homodromic cycles. Grishina and Buch later proposed another algorithm where the initial structure does not satisfy the ice rules and contains many point defects with excess inbound or outbound hydrogen bonds, called D- and L-type Bjerrum defects.<sup>13</sup> This algorithm randomly chooses a hydrogen bond at a defect and moves the defect by reversing the orientation of the bond. The defect annihilates when it encounters another defect of another type. This abacus-like action is repeated until all defects are erased. Both algorithms are stochastic and time-consuming.

Ion-doped ice, which involves replacing some water molecules with ions, has gained attention recently. Normally, ice does not form solid solutions. When salt water freezes, only water freezes, and brine rejection occurs.  $\text{NH}_4\text{F}$  is an exception. Neither the cation nor the anion enters the interstitial space but replaces a water molecule, following a different “ionic rule”: the  $\text{NH}_4^+$  cation donates four hydrogen bonds to neighboring water molecules, and the  $\text{F}^-$  anion accepts four hydrogen bonds. To create such ice for molecular simulations, some water molecules need to be replaced with ions. This requires a global reorientation of the hydrogen bonds.  $\text{NH}_4\text{F}$  is called the “hydrogen-disorder agent” because it interferes with the ice rules and inhibits the formation of hydrogen-ordered ice and crystalline phases with odd-numbered rings.<sup>14,15</sup> In reality,  $\text{NH}_4\text{F}$  also reduces the disorder in hydrogen-disordered ice by inhibiting the free orientation of water molecules.

There is a similar problem with semiclathrate hydrates. In tetrabutylammonium bromide (TBAB) hydrates, the  $\text{TBA}^+$  cation and the  $\text{Br}^-$  anion each replace one or two water molecules and enter the hydrogen bonding network of the clathrate hydrate.<sup>16</sup> To simulate these solid solutions on a computer, it is essential to create hydrogen-disordered networks that follow the ionic and ice rules.

The commonly used stochastic algorithms are not suitable for simulating ice crystals that are larger than tens of thousands of molecules.<sup>12,13</sup> As a solution, a new algorithm called *GenIce* has been developed based on graph theory. This algorithm can create hydrogen-disordered structures with a computation time that increases only linearly with the number of molecules.<sup>17</sup> Using *GenIce*, it has become possible to generate many ice structures of over  $1 \times 10^6$  molecules and investigate unique properties of ice that become apparent at such huge sizes.<sup>10</sup> *GenIce* has been used in various studies, including *ab initio* prediction of thermodynamic properties,<sup>18</sup> ice nucleation,<sup>19</sup> functionalities of the antifreeze proteins,<sup>20</sup> exploration of the phase diagram of water,<sup>21,22</sup> gas hydrates and their inhibitors,<sup>23,24</sup>

high-pressure ices,<sup>25,26</sup> detection of local order in ice,<sup>27,28</sup> and interfaces.<sup>29</sup>

We have made enhancements to the *GenIce* algorithm in this paper. Our new algorithm, called *GenIce-core*, has the capability to generate uniformly random molecular orientations. It can create ice surface structures where not all water molecules have four neighbors and doped ice structures that obey the ionic rule.

The previous version of *GenIce* software included all the steps involved in creating ice structures, such as generating crystal structures, applying the ice rules, depolarization, molecular arrangement, and output in different formats. However, *GenIce-core* only contains the core algorithm for the ice rules and depolarization. It is designed as an API that can be easily integrated into other software for crystal design purposes.

## ALGORITHM

First, the algorithm for pure ice, which may have surfaces and/or Bjerrum defects, is described in this section. It is then expanded to address systems containing doped ions. Here, we describe the algorithm for a simplified two-dimensional ice model, which can be applied to realistic three-dimensional ice and even hypothetical ices in higher dimensions using the same procedure.

### Pure ice

- 1a. An undirected graph  $G$  [Fig. 1(a)] representing the skeletal structure of a hydrogen bond network of ice is given. The degree of  $G$  must be at most 4.
- 1b. Water molecules at the ice surface or at the sites adjacent to a Bjerrum defect have less than four neighboring molecules [Fig. 1(b)]. We add virtual nodes and corresponding edges so that every node has four neighbors. In Fig. 1(b), the virtual nodes are represented by yellow dots.
- 1c. We split each normal node into two nodes. One of the nodes [e.g., blue nodes in Fig. 1(c)] will have two randomly chosen edges out of the four, while the other (pink nodes) will have the remaining two edges. No direct connection exists between the two nodes. This process results in a graph called  $H$ , which consists only of cycles and chains since it has a maximum degree of two. Virtual nodes remain unsplit. We refer to this process as *noodlization*, as shown in Fig. 1(c).
- 1d. The paths along each cycle and chain prepared in step 1c can be long, so they often have self-intersections. These self-intersections happen when there are lattice sites at which both split nodes have two edges of the same color, as in Fig. 1(c). Each path is further decomposed into chains and cycles to eliminate these self-intersections. For example, the light green path in Fig. 1(c) is decomposed into one chain (solid) and one cycle (dotted) in Fig. 1(d).
- 1e. To give cycles and chains direction, we need to determine the orientation of the hydrogen bonds. This orientation determines the net dipole moment of the system. To minimize total polarization, we assign directions randomly to the chains

and cycles [Fig. 1(e)]. Finding the optimal combination of path directions is time-consuming, as it involves a knapsack problem. Therefore, we use a finite-step Monte Carlo method to find an approximate solution. It is worth noting that compact cycles (cycles that do not traverse the simulation cell under periodic boundary conditions) have no net dipole moments. Hence, we can assign their directions freely.

- 1f. A hydrogen-disordered ice structure is generated after removing the virtual nodes and edges and merging the split nodes again. In the ice structure, all hydrogen bonds are arranged randomly within the constraints of the ice rules, and the net dipole moment is minimal [see Fig. 1(f)].
- 1g. Water molecules are placed at the node positions so that the hydrogen atoms direct along the arrows in this hydrogen bonding network [Fig. 1(g)]. This step is not included in GenIce-core.

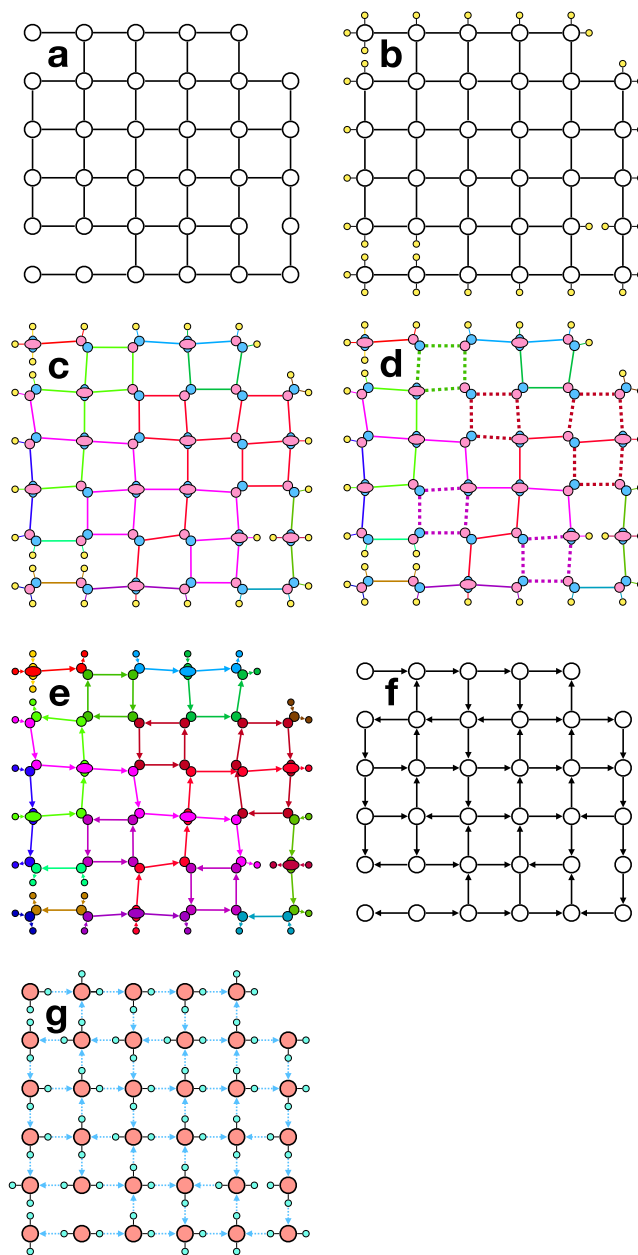
This is an  $O(N)$  algorithm and is as fast as the previous GenIce algorithm.<sup>17</sup> Each step of the algorithm is suitable for parallel execution. The processing time and the randomness of the obtained digraph are discussed in later sections.

### Doped ice

Ions that replace water molecules in ice do not follow the ice rules. The hydrogen bonds at these ions cannot be rearranged, which means that if they are included in the chain in Fig. 1(e), the chain cannot be used to adjust the net dipole moment by reversing its direction.

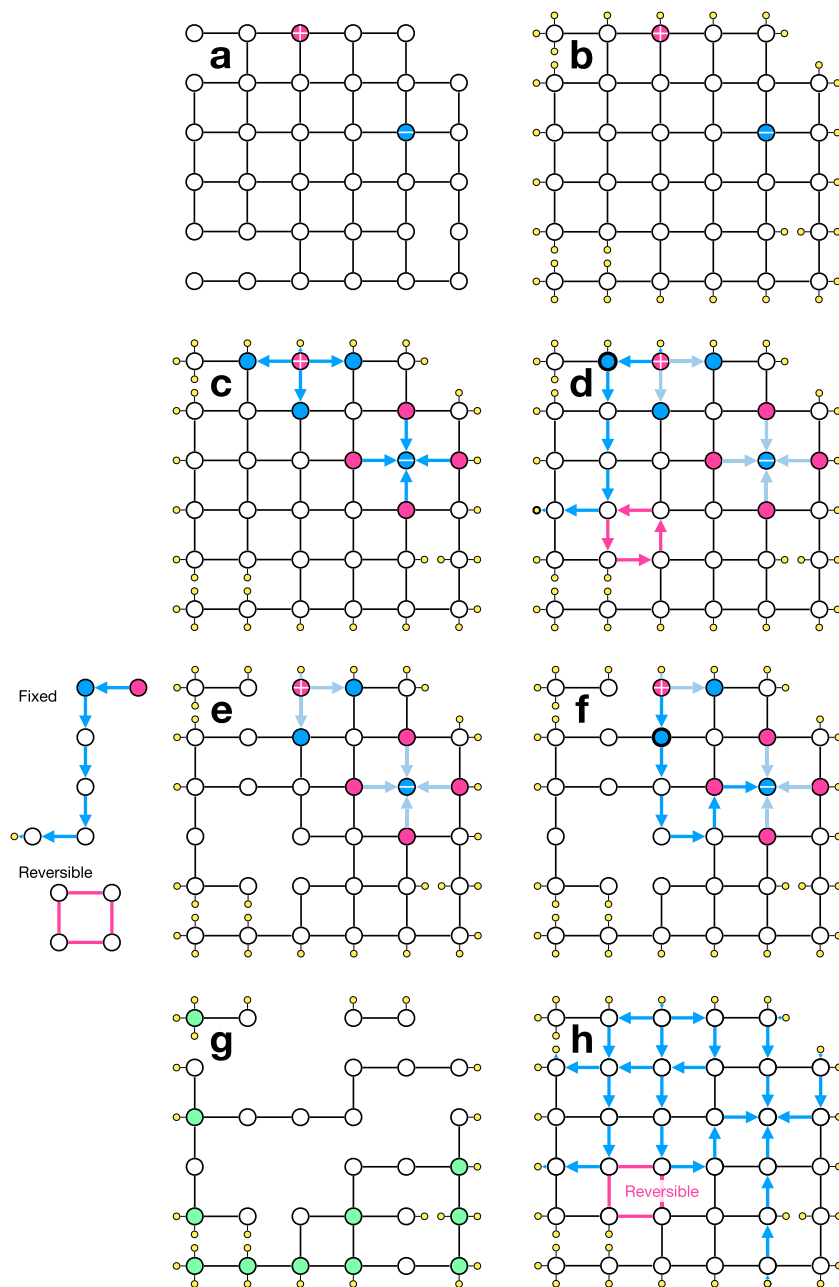
Here is an example of an ice cluster being one  $F^-$  anion and one  $NH_4^+$  cation doped. These ions do not enter interstitially but replace water molecules to form a solid solution. They have four outward edges for  $NH_4^+$  and four inward edges for  $F^-$ ; their directions cannot be changed. The edges whose directions are fixed are preprocessed accordingly.

- 2a. We identify the location of dopants in an undirected graph  $G$  [as shown in Fig. 2(a)].
- 2b. The virtual nodes and edges are added as in step 1b so that all nodes become 4-degree [Fig. 2(b)].
- 2c. The normal nodes adjacent to the ions, colored blue and pink in the figure, are put in a queue [Fig. 2(c)].
- 2d. Start a random walk from one of the nodes in the queue. The selected node is removed from the queue. The random walk is allowed to pass an edge once. Therefore, it must end at a yellow node [Fig. 2(d)]. The random walk also ends when it connects a blue node with a pink node.
- 2e. The edges that are part of the path are removed from  $G$  [as shown in Fig. 2(e)]. When the path has self-intersections, it is divided into one or more cycles and one open path, as in the case of step 1d. The direction of the open path is not flexible because it has one or two nodes connected to the ions. On the other hand, there are two possible directions for each cycle. The open path and cycles are recorded separately in memory.
- 2f. Step 2d is repeated until the queue becomes empty [Fig. 2(f)]. A remaining undirected graph with only nodes of even degree is eventually obtained [Fig. 2(g)].



**FIG. 1.** The GenIce-core algorithm is illustrated. (a) Original skeletal graph of ice with surfaces. (b) Addition of virtual nodes. (c) Noodlization, i.e., each 4-degree node is divided into two 2-degree nodes. (d) Removal of self-intersections. (e) Conversion from the undirected graph to a directed graph. (f) Recombination of the split nodes. (g) Conversion from the directed graph to the corresponding molecular configuration.

- 2g. For nodes of degree 4, apply the *noodlization* given in step 1c. It is not necessary to split the nodes of degree 2 [Fig. 2(g)].
- 2h. Apply step 1d and the following procedure.

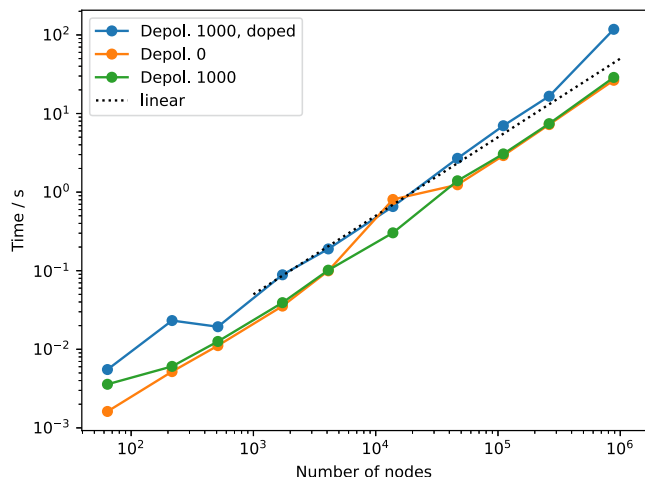


**FIG. 2.** The preprocess for doped ice is illustrated. (a) Original undirected graph with a pair of ions doped. (b) Addition of virtual nodes. (c) Directions of the four edges connecting the ions and surrounding nodes. (d) A random walk starting from a blue node. (e) Decomposition of the path. Open and closed paths are recorded separately. (f) Another random walk. (g) Resultant graph without the ions. Green nodes with four edges are subject to the split in step 1c. (h) Final graph shown together with the walking paths.

- 2i. In step 1e, the derived cycles obtained in step 2e are subject to depolarization.

Steps 2a–2e are similar to those implemented in the previous GenIce algorithm.<sup>17</sup> The algorithm processes all edges

only once, making it an  $O(N)$  algorithm. If there are many fixed edges at the beginning, the search for a self-avoiding path may encounter dead ends and fail. In such a case, the search must be repeated, which requires additional computation time.



**FIG. 3.** Process time against the number of water molecules. Each point on the graph represents the average time from ten trials. The orange and green symbols represent the time it takes to generate ice structures without depolarization and with 1000 Monte Carlo iterations for depolarization, respectively. The blue symbols show the time it takes to create ice with a pair of ions doped at lattice points that are far apart. The dotted line is a linear function of the number of molecules in the eye guide. These benchmark tests were performed on a MacBook Air M1.

## IMPLEMENTATION AND EVALUATIONS

### API

The algorithm is implemented in pure Python, available as a Python Package Index (PyPI) module, and easily installed with a single command:

```
pip install genice-core
```

The application programming interface (API) comprises only one function called `ice_graph()`. This function requires users to prepare an undirected graph representing the skeletal network structure of an ice crystal. If the coordinates of the nodes are also provided, depolarized structures can be produced. The function generates a directed graph satisfying both the ionic and ice rules. For further details on how to use the function, please refer to the GitHub repository.

### Benchmark tests

In order to evaluate the execution speed, hydrogen-disordered ice of various sizes was generated under periodic boundary conditions. Figure 3 illustrates the computation time for pure ice with and without depolarization and ice with a doped ion pair. The computation time increases linearly with the system size when no dopant is present. It takes slightly longer to generate ice with dopant, but it is still practical time for generating a system with  $10^6$  molecules.

The GenIce-core repository includes the code that was used for this benchmark.

**TABLE I.** Probability of the directional pattern of the six hydrogen bonds present along a six-membered ring in hydrogen-disordered ice I. The values for ice consisting of  $128^3 =$  two million water molecules and four million rings generated by the GenIce-core algorithm are compared with theoretical values based on a Bethe tree approximation. All inversions and sequences with shifted starting points are considered to be identical, i.e.,  $\rightarrow\rightarrow\rightarrow\rightarrow\leftarrow\leftarrow$  and  $\rightarrow\leftarrow\leftarrow\leftarrow\rightarrow\rightarrow$  are counted as identical. The last column shows the case of proton-ordered ice II for comparison.

Pattern	Approx.	GenIce-core	Ice II
$\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow$	$64/365 = 0.175\ 34$	0.174 66	0.142 86
$\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\leftarrow$	$96/365 = 0.263\ 01$	0.265 19	0
$\rightarrow\rightarrow\rightarrow\rightarrow\leftarrow\leftarrow$	$96/365 = 0.263\ 01$	0.262 44	0
$\rightarrow\rightarrow\rightarrow\leftarrow\leftarrow\leftarrow$	$24/365 = 0.065\ 75$	0.063 51	0
$\rightarrow\rightarrow\rightarrow\leftarrow\leftarrow\leftarrow$	$48/365 = 0.131\ 51$	0.125 42	0.428 57
$\rightarrow\rightarrow\leftarrow\rightarrow\leftarrow\leftarrow$	$12/365 = 0.032\ 88$	0.036 59	0.428 57
$\rightarrow\rightarrow\leftarrow\rightarrow\leftarrow\leftarrow$	$24/365 = 0.065\ 75$	0.069 84	0
$\rightarrow\leftarrow\leftarrow\rightarrow\leftarrow\leftarrow$	$1/365 = 0.002\ 74$	0.002 36	0

### Homogeneity tests

GenIce has a feature that checks whether the generated structure is uniformly random.<sup>30</sup> This feature can be exemplified with hexagonal ice I, a type of hydrogen-disordered ice where the hydrogen bonding network consists exclusively of six-membered rings. Each hydrogen bond along a ring is oriented randomly, and when we examine their directions along a cycle, the probability of finding a specific sequence of forward and backward edges should conform to a statistical rule. An approximated probability value can be obtained for each sequence, as explained in the Appendix. Table I demonstrates that the hydrogen bond networks of ice Ih produced by GenIce-core are sufficiently random.

### SUMMARY

We have developed a new algorithm that can quickly generate hydrogen-disordered ice structures. This algorithm is suitable for ice surfaces and ice doped with some ions. Additionally, an API is provided for easy integration into other crystal structure design software. The source code is available on GitHub at the following repository: <https://github.com/genice-dev/genice-core>. This algorithm will also be integrated into the upcoming GenIce package version 2.2.<sup>30</sup>

### ACKNOWLEDGMENTS

The original idea of the noodlization was provided by Professor Tadashi Sakuma. The present work was supported by JSPS KAKENHI Grant No. 21H01047 and the Research Center for Computational Science in providing computational resources (Project No. 23-IMS-C028).

### AUTHOR DECLARATIONS

#### Conflict of Interest

The authors have no conflicts to disclose.



## Author Contributions

**Masakazu Matsumoto:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Takuma Yagasaki:** Investigation (equal); Writing – review & editing (equal). **Hideki Tanaka:** Investigation (equal); Writing – review & editing (equal).

## DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## APPENDIX: BETHE TREE APPROXIMATION FOR THE DIRECTIONAL PATTERNS IN A CYCLE

In Table I, there are eight different hydrogen bond patterns in a six-membered cycle. To simplify, the right and left arrows are labeled as 0 and 1. These patterns are identified as 000000, 000001, 000011, 000101, 000111, 001001, 001011, and 010101.

### Pattern 000000

A cycle of the pattern 000000 is called a homodromic cycle.<sup>31,32</sup> In this configuration, each node on the cycle has one inbound edge and one outbound edge. Because each water molecule should have four hydrogen bonds, one of the two remaining edges must be inbound and the other must be outbound, resulting in two possibilities. We call this kind of node C. Its symmetry number,  $\sigma$ , is two (000000 or 111111) and, therefore, the number of possible configurations for nodes on this type of cycle is  $2^6\sigma = 128$ .

### Pattern 000001

The sequence 01 (literally  $\rightarrow\leftarrow$ ) indicates that two adjacent edges are in opposite directions, and the central node has two inbound hydrogen bonds. Due to the ice rules, two other hydrogen bonds must be outbound, so there is no choice for these edges. We call this kind of node NC. The symmetry number for this sequence is 12 (100000, 010000, ..., 000001 and their inverses, i.e., 011111,

101111, ..., 111110). The number of possible configurations is  $2^4\sigma = 192$ .

### Pattern 000011 and the others

Similarly, we can calculate the number of possible configurations based on the symmetry number and the number of type-C nodes in the cycle. All possible configurations are listed in Table II.

The calculations mentioned above are applicable to cycles of sizes other than six-membered ones. This is a Bethe tree approximation that ignores the correlation between edges through the other part of the graph where the six-membered cycle is embedded.

## REFERENCES

- V. F. Petrenko and R. W. Whitworth, *Physics of Ice* (Oxford University Press, Oxford, 1999).
- T. M. Gasser, A. V. Thoeny, A. D. Fortes, and T. Loerting, "Structural characterization of ice XIX as the second polymorph related to ice VI," *Nat. Commun.* **12**(1), 1128 (2021).
- H. Tanaka, M. Matsumoto, and T. Yagasaki, "On the role of intermolecular vibrational motions for ice polymorphs. IV. Anisotropy in the thermal expansivity and the nonaffine deformation for ice IX and III," *J. Chem. Phys.* **157**(17), 174505 (2022).
- W. F. Giauque and J. W. Stout, "The entropy of water and the third law of thermodynamics. The heat capacity of ice from 15 to 273 K," *J. Am. Chem. Soc.* **58**(7), 1144–1150 (1936).
- J. D. Bernal and R. H. Fowler, "A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions," *J. Chem. Phys.* **1**(8), 515–548 (1933).
- T. Hondoh, K. Azuma, and A. Higashi, "Self-interstitials in ice," *J. Phys. Colloq.* **48**(C1), C1-183–C1-187 (1987).
- C. Kobayashi, S. Saito, and I. Ohmine, "Mechanism of fast proton transfer in ice: Potential energy surface and reaction coordinate analyses," *J. Chem. Phys.* **113**, 9090 (2000).
- H. Gränicher, C. Jaccard, P. Scherrer, and A. Steinemann, "Dielectric relaxation and the electrical conductivity of ice crystals," *Discuss. Faraday Soc.* **23**(0), 50–62 (1957).
- L. Pauling, "The structure and entropy of ice and of other crystals with some randomness of atomic arrangement," *J. Am. Chem. Soc.* **57**(12), 2680–2684 (1935).
- M. Matsumoto, T. Yagasaki, and H. Tanaka, "On the anomalous homogeneity of hydrogen-disordered ice and its origin," *J. Chem. Phys.* **155**(16), 164502 (2021).
- J. Lasave, S. Koval, A. Laio, and E. Tosatti, "Proton strings and rings in atypical nucleation of ferroelectricity in ice," *Proc. Natl. Acad. Sci. U. S. A.* **118**(1), e2018837118 (2021).
- A. Rahman and F. H. Stillinger, "Hydrogen-bond patterns in liquid water," *J. Am. Chem. Soc.* **95**(24), 7943–7948 (1973).
- N. Grishina and V. Buch, "Structure and dynamics of orientational defects in ice I," *J. Chem. Phys.* **120**(11), 5217–5225 (2004).
- C. G. Salzmann, Z. Sharif, C. L. Bull, S. T. Bramwell, A. Rosu-Finsen, and N. P. Funnell, "Ammonium fluoride as a hydrogen-disordering agent for ice," *J. Phys. Chem. C* **123**(26), 16486–16492 (2019).
- Z. Sharif, J. J. Shephard, B. Slater, C. L. Bull, M. Hart, and C. G. Salzmann, "Effect of ammonium fluoride doping on the ice III to ice IX phase transition," *J. Chem. Phys.* **154**(11), 114502 (2021).
- W. Shimada, M. Shiro, H. Kondo, S. Takeya, H. Oyama, T. Ebinuma, and H. Narita, "Tetra-*n*-butylammonium bromide-water (1/38)," *Acta Crystallogr. C* **61**(2), o65–o66 (2005).
- M. Matsumoto, T. Yagasaki, and H. Tanaka, "Novel algorithm to generate hydrogen-disordered ice structures," *J. Chem. Inf. Model.* **61**(6), 2542–2546 (2021).

**TABLE II.** All possible configurations for a six-membered cycle and the probabilities of finding them.

Label	No. of C	Symmetry no.	No. of conf.	Probability
000000	6	2	$2 \times 2^6 = 128$	128/730
000001	4	12	$12 \times 2^4 = 192$	192/730
000011	4	12	$12 \times 2^4 = 192$	192/730
000101	2	12	$12 \times 2^2 = 48$	48/730
000111	4	6	$6 \times 2^4 = 96$	96/730
001001	2	6	$6 \times 2^2 = 24$	24/730
001011	2	12	$12 \times 2^2 = 48$	48/730
010101	0	2	$2 \times 2^0 = 2$	2/730
Total			730	1

- <sup>18</sup>B. Cheng, E. A. Engel, J. Behler, C. Dellago, and M. Ceriotti, "Ab initio thermodynamics of liquid and solid water," *Proc. Natl. Acad. Sci. U. S. A.* **116**(4), 1110–1115 (2019).
- <sup>19</sup>P. M. Piaggi, J. Weis, A. Z. Panagiotopoulos, P. G. Debenedetti, and R. Car, "Homogeneous ice nucleation in an ab initio machine-learning model of water," *Proc. Natl. Acad. Sci. U. S. A.* **119**(33), e2207294119 (2022).
- <sup>20</sup>K. Mochizuki and V. Molinero, "Antifreeze glycoproteins bind reversibly to ice via hydrophobic groups," *J. Am. Chem. Soc.* **140**(14), 4803–4811 (2018).
- <sup>21</sup>S. L. Bore, P. M. Piaggi, R. Car, and F. Paesani, "Phase diagram of the TIP4P/Ice water model by enhanced sampling simulations," *J. Chem. Phys.* **157**, 054504 (2022).
- <sup>22</sup>T. Matsui, T. Yagasaki, M. Matsumoto, and H. Tanaka, "Phase diagram of ice polymorphs under negative pressure considering the limits of mechanical stability," *J. Chem. Phys.* **150**(4), 041102 (2019).
- <sup>23</sup>T. Yagasaki, M. Matsumoto, and H. Tanaka, "Adsorption of kinetic hydrate inhibitors on growing surfaces: A molecular dynamics study," *J. Phys. Chem. B* **122**(13), 3396–3406 (2018).
- <sup>24</sup>Y. Lee, W. Go, Y. Kim, J. Lim, W. Choi, and Y. Seo, "Molecular guest exchange and subsequent structural transformation in CH<sub>4</sub>-CO<sub>2</sub> replacement occurring in sH hydrates as revealed by <sup>13</sup>C NMR spectroscopy and molecular dynamic simulations," *Chem. Eng. J.* **455**(2), 140937 (2023).
- <sup>25</sup>A. Toffano, J. Russo, M. Rescigno, U. Ranieri, L. E. Bove, and F. Martelli, "Temperature- and pressure-dependence of the hydrogen bond network in plastic ice VII," *J. Chem. Phys.* **157**(9), 094502 (2022).
- <sup>26</sup>X. Zhang, Y. Yao, H. Li, A. Python, and K. Mochizuki, "Fast crystal growth of ice VII owing to the decoupling of translational and rotational ordering," *Commun. Phys.* **6**(1), 164 (2023).
- <sup>27</sup>H. Doi, K. Z. Takahashi, and T. Aoyagi, "Searching local order parameters to classify water structures of ice Ih, Ic, and liquid," *J. Chem. Phys.* **154**(16), 164505 (2021).
- <sup>28</sup>Q. Kim, J.-H. Ko, S. Kim, and W. Jhe, "GCIceNet: A graph convolutional network for accurate classification of water phases," *Phys. Chem. Chem. Phys.* **22**(45), 26340–26350 (2020).
- <sup>29</sup>S. Mukherjee and B. Bagchi, "Entropic origin of the attenuated width of the ice–water interface," *J. Phys. Chem. C* **124**(13), 7334–7340 (2020).
- <sup>30</sup>M. Matsumoto, T. Yagasaki, and H. Tanaka, "GenIce: Hydrogen-disordered ice generator," *J. Comput. Chem.* **39**(1), 61–64 (2018).
- <sup>31</sup>W. Saenger, "Circular hydrogen bonds," *Nature* **279**(5711), 343–344 (1979).
- <sup>32</sup>T. Nakamura, M. Matsumoto, T. Yagasaki, and H. Tanaka, "Thermodynamic stability of ice II and its hydrogen-disordered counterpart: Role of zero-point energy," *J. Phys. Chem. B* **120**(8), 1843–1848 (2016).